

## DOCUMENT RESUME

ED 217 485

CS 207 043

AUTHOR Hoetker, James; And Others  
TITLE Creating Essay Examination Topics.  
PUB DATE [82]  
NOTE 15p.; Best copy available.

EDRS PRICE MF01/PC01 Plus Postage.  
DESCRIPTORS \*Essay Tests; Higher Education; Minimum Competency Testing; Program Development; \*Research Methodology; State Programs; \*Teacher Certification; Teacher Education; \*Writing (Composition); Writing Instruction; \*Writing Research

IDENTIFIERS \*Essay Topics; Florida; \*Teacher Competencies

## ABSTRACT

The efforts of researchers to produce a pool of approximately 30 validated and field-tested topics for use in future administrations of the Florida Teacher Certification Examination in writing are reviewed in this paper. Various sections of the paper (1) describe the procedures followed; (2) explain the reasons for making certain important decisions, including the phrasing of topics and selection of criteria for choosing certain topics and for rejecting others; and (3) present the results of the field tests of the chosen topics. Based on the findings of the field tests, the paper concludes that responsible researchers and testwriters can no longer take essay examination topics for granted. (FL)

\*\*\*\*\*  
\* Reproductions supplied by EDRS are the best that can be made \*  
\* from the original document. \*  
\*\*\*\*\*

ED217485

U.S. DEPARTMENT OF EDUCATION  
NATIONAL INSTITUTE OF EDUCATION  
EDUCATIONAL RESOURCES INFORMATION  
CENTER (ERIC)

☐ This document has been reproduced as  
received from the person or organization  
originating it.  
☒ Minor changes have been made to improve  
reproduction quality.

• Points of view or opinions stated in this docu-  
ment do not necessarily represent official NIE  
position or policy.

CREATING ESSAY EXAMINATION TOPICS

James Hoetker, Gordon Brossell, and Barbara E. Ash

English Education

209 Education

Florida State University

Tallahassee, Florida 32306

(904) 644-6553

CS227043

BEST AVAILABLE COPY

## CREATING ESSAY EXAMINATION TOPICS

In the autumn of 1980 the Florida Department of Education asked us to provide a pool of about 30 validated and field-tested topics for use in future administrations of the Teacher Certification Examination. At each step in the process of creating these topics, we were guided by the results of available research. We wish to describe the topic development process we employed, to explain the reason for some of our decisions, and to report some discoveries we made along the way. We cannot, of course, quote any of the topics that were developed, since circulation of the topics would jeopardize the testing program. Instead, an example is given that resembles an actual topic in its important features. Similarly, the sample examination instructions are intended only to illustrate the approach taken.

We will first describe and comment on the decisions that were made about time, options, mode of discourse, and specification of rhetorical context. Then we will outline the procedures followed in producing and testing the topics. Finally, we will report the field test results.

Time and the Provision of Options

The test administrators had allotted 45 minutes for the writing examination in the first administrations of the test. We used the same time period in the field tests. Later, on the basis of the literature we could cite and on the basis of reports from field test participants, we urged that the writing examination be extended to a full hour. Our

recommendation was accepted.

The Council on Teacher Education had decided that students taking the writing examination would be given two topics to choose from. Providing options seemed fairer and more defensible than requiring students to write on a single set topic. Only two topics were provided, though, since giving three or more would probably tempt students to waste valuable time making decisions. In the field test, we took the opportunity to compare the scores of student essays written under both single-topic and optional-topic conditions.

#### Mode of Discourse

It seemed to us that, from the students' point of view, the most crucial modal distinction would be that between a personal essay and a "scholarly" essay. Educated persons can reasonably be expected to write competent accounts of their own feelings, opinions, or experiences. But not everyone has been trained up to the same level in the conventions of objective analysis and reporting. Calling for a third-person essay would mean increasing the already considerable advantages of students majoring in language-centered disciplines.

We phrased both the examination instructions and the topics themselves to call for informal personal statements about the subject matter identified in the topics. The great majority of students had no trouble interpreting the topics as we intended and responding appropriately. A few, though, ignored the instructions and tried (unsuccessfully) to

impress, the readers with an apodictic impersonality. These students probably allowed memories of their school writing experiences to dictate their misreading. For them, the situational cues—This is a test!—were so powerful as to cause them to ignore or disbelieve repeated requests to write a personal essay.

At any rate, the readers were instructed not to penalize writers for ignoring or misreading such directions, but to apply the rating criteria in the customary manner to whatever species of prose a student had produced.

#### Specification of Rhetorical Context

A distinction needs to be made between necessary rhetorical context and excessive rhetorical context. Brossell, in the preceding article, presented evidence that too much information can be counterproductive. Students certainly are aided by some contextual information, but there is a point at which the sheer number of the topic-writer's words overburdens students and interferes with their composing behaviors. It cannot yet be said where the point is, or whether it differs according to student characteristics or writing situations.

In our attempt to find a balance between the extremes of inadequate and excessive specification, we introduced information about audience into the examination instructions and allowed the topic proper to carry information about the desired voice. We felt the examination situation itself adequately conveyed information about form and purpose (at least

it did when reinforced by such indirect cues as the use of the word essay). As noted in the preceding sections, this amount of contextual information provided sufficient guidance to the great majority of students.

#### Instructions about Audience

We thought it unnecessary for our purposes to ask students to assume any other role than students-writing-an-examination. That role is certainly real enough to be motivating. So, rather than invent some imaginary audience to write for, we described for the students the actual readers and how they would be reading. The audience description used in the field test read something like this:

Your essay will be read by three experienced English teachers who will be judging its quality as a prose composition. They will be fully aware that you wrote under pressure and without preparation. They will not be concerned about whether your opinions agree with theirs. The only issue is the quality of your writing—how well you say whatever you choose to say.

#### The Phrasing of the Topics

We had concluded that topics carrying a moderate information load tended to produce more focussed and well-organized essays than topics that were either exceedingly terse or overly-detailed. We wanted to

produce topics at a moderate information level that shared important grammatical and syntactic features, so as to limit the range of possible misreadings across topics. We at length decided on a two-sentence format in which all topics would be cast.

SENTENCE 1. A declarative sentence (a simple sentence where possible) which makes a generalization about a particular subject matter.

SENTENCE 2. An interrogatory sentence (also simple where possible) which asks the student to make a personal response to an instance of the subject matter identified in the first sentence.

Here is a sample topic made up for purposes of illustration:

TOPIC: Everyone at some time suffers the loss of some especially valued person or thing. How have you or someone you know coped with such a tragic loss?

We think that the phrasing of the second sentence as a direct question is important; and we have a strong hunch that such a question is more helpful to the students than an alternative phrasing such as, "Explain your own reactions to such a situation." In reading essays produced earlier in the course of our work, we had observed that the topics which happened to end with a direct question tended to produce better organized essays. Students were quite often able to take a question and turn it around—transform it—into a topic sentence which suggested its own pattern of development.

Presented with a question in one of these forms,

"How have you \_\_\_\_\_-ed with \_\_\_\_\_?"

"Why do you think \_\_\_\_\_ is \_\_\_\_\_?"

students often converted the question into a statement in these forms,

"I have \_\_\_\_\_-ed with \_\_\_\_\_ by \_\_\_\_\_-ing."

"I think \_\_\_\_\_ is \_\_\_\_\_ because \_\_\_\_\_."

For example, asked the question in the sample topic, a student might begin her essay by writing: "I coped with the loss of my dog, a Labrador I'd had since I was nine, by simply falling apart." Already, without any of the threshing around typical of opening paragraphs, the student has found her subject--an explanation and illustration of what she means by "falling apart." The details, after this start, will follow easily enough, probably in chronological order.

By contrast, the same topic with its second sentence cast in the form,

"Explain how you have \_\_\_\_\_-ed with \_\_\_\_\_."

does not (for reasons we do not yet understand) seem to hold the same potential for directing students to a starting place. Even less helpful, we would guess, are topics ending with a tacked-on "Why?"

Our feelings about the importance of the question phrasing were strong enough to guide our topic writing. Unlike most hunches about the teaching and testing of writing, this one has the virtue of being specific enough to be tested in a straight-forward experimental comparison.

Procedures for Developing Topics

We began by calling in five experienced English teachers and giving them the topic criteria. After an orientation and some preliminary brainstorming, they were asked to submit within a few days at least 20 topic ideas apiece—not fully phrased topics, just suggestions of appropriate subjects. They gave us over 200 topic ideas. From those we chose 50 promising ones and phrased them all in the two-sentence format discussed above.

The 50 topics were reviewed and criticized in writing by two English professors with long experience directing writing programs. Then the topics were similarly reviewed by three undergraduate education students (who, in fact, found some problems that the older reviewers were not alert to). Thirty-three topics survived this review process, and they were edited one final time and field-tested.

A total of 350 education students at three Florida universities (two public and one private) wrote essays on the examination topics and then completed an accompanying questionnaire.

Preparation of Examination Forms

We prepared the examination packages in such a way that we could both compare student performances on one-topic and two-topics examination forms and get a reading on the "popularity" of each topic. The topics were arbitrarily assigned numbers. Then odd numbers were laid

out against even numbers to produce a gridwork of 272 cells, which was reduced to 250 cells by cancelling two eleven-cell diagonals. The grid was used by a typist as a guide as she typed topics on the two-topics examination sheets. On successive copies of the examination, for example, topic 2 would be paired with topics 1, 3, 5, 7 and so forth through 33. Then topics 4, 6, 8 and so forth would be paired with from fourteen to sixteen odd-numbered topics. Another set of examination sheets was prepared on which only a single topic appeared.

The single-topic form of the examination was given to 123 students, while 227 were given the two-topics form. The mean score on the two-topics form was slightly and non-significantly higher: 6.50 as opposed to 6.40. The evidence about the effects of optional topics remains so mixed and uncertain, that we can only agree with COTE that the optional topic is preferable because it gives an appearance of fairness.

#### Other Field Test Results

##### "Popularity" of Topics

Since each topic had been paired numerous times with other topics, it was possible to determine "popularity" by calculating how often a topic had actually been chosen by students, as a percentage of the number of times it might have been chosen. Some topics were clearly more appealing to students than others. One topic was chosen fourteen times out of fourteen chances, while another was chosen only three times

out of sixteen. But except in two extreme cases, popularity was quite unrelated to the mean scores of essays written on the topics.

We had also asked students on the accompanying questionnaire to rate the topics as to how difficult, how interesting, and how appropriate they were. In the cases of all but three topics that were later rejected, these ratings were also unrelated to scores.

### Personal and Public Topics

Students who were given the two-topics form of the examination were asked why they had rejected the topic on which they did not choose to write. Some students explained they had found the rejected topics "too personal to write about." These students often went on to explain that they liked a controversial topic they could "take a stand on." Others, in about equal numbers, had rejected argumentative topics ("I don't like to argue") in favor of topics that allowed them to write about what they really knew--themselves.

We had, as explained earlier, taken the importance of mode of discourse into account when we decided to ask students to write subjectively. But the student comments made us realize we had ignored another equally important dimension of the matter: the source of the essay's content. We had not taken into account that, even if all students are to write personal essays, it makes a difference whether the topic allows students to give their opinions on a public issue or demands that they introspect and report on their inner lives.

When we re-examined our topics we found, first, that each topic

could clearly be assigned to either a "public" or "private" category, and, second, that we had ~~by~~ some pleasant chance ~~written~~ almost equal numbers of the two sorts of topics.

When we compared mean scores of essays on the public and private topics, we found them to be almost identical—~~6.38~~ 6.38 and 6.39. These scores were not significantly different from the scores on essays written by students who were not given options. Still, student preferences were expressed strongly and often. So, again aiming to maximize the fairness of the whole proceeding, we recommended that student preferences be taken into account by choosing one public and one private topic as the options for each form of the examination.

#### Distribution of Topic Scores

The mean scores for the essays on the 30 topics retained after the field tests ranged from 5.5 to 8.1, with a mean of 6.4. We found it impossible, though, to interpret even the extreme scores as indicators of topic difficulty or quality. There were three complicating factors. First, only the two or three scores at the extremes of the distribution were significantly different from the mean, and very few students had written on the topics with extreme scores.<sup>1</sup> Second, standard deviations for the mean scores of all topics were extremely large. Essays of high, average, and low quality were written to all topics. Third, and most seriously, there was evidence of a confounding of topic and writer effects that we had no way to untangle, lacking independent verbal ability scores on the students. The mean score, for example, of essays

written by fifteen English Education majors on ten topics was 8.2—almost two points above the mean. The mean score of the 40 essays written by private university students was 7.5. The mean scores for the topics these students happened to write on were, of course, elevated. It is reasonable to suppose there were similar sub-groups of students with lower verbal skills who would tend to lower the mean score of topics they wrote on.

Not being able to interpret mean score differences with any confidence, we submitted to the DOE all the topics that were satisfactory by the other criteria, cautioning that future score distributions could not necessarily be predicted from field test distributions, and that mean scores should not be used as a basis for labelling topics as harder or easier than others.

#### The Rejected Topics

The field tests identified three bad topics that had survived the editing processes. Students scored poorly on them (a mean of 5.2), chose them seldom, and found them difficult, uninteresting, inappropriate, and requiring special knowledge. Two of the topics—one about American neglect of the urban environment and another about "favorite gadgets"—were just mistakes. No one had anything of interest to say about either.

The third rejected topic was a much more instructive case, however. It gave us our only clear instance of a direct connection between the

subject matter of a topic and the style of essays written on it. This topic called on students to describe their dream home. A very similar exercise had been used by one of us for years in writing classes, and it had always given excellent results when students had several days to work out their ideas. But every essay written on this topic in 45 minutes was utterly disorganized. Apparently, the students had fallen into daydreaming on paper, putting down each new idea as it occurred in stream-of-consciousness fashion. That these students could not write well on this topic, which had been successfully used in writing classes, is one more reminder that testing writing is not the same as teaching writing.

#### Conclusion

These three studies establish that responsible researchers and test-writers cannot, in the future, take essay examination topics for granted.

FOOTNOTE

<sup>1</sup> The design feature which enabled us to compare scores on one-topic and two-topics forms of the test also, unfortunately, prevented us from ensuring that approximately equal numbers of students wrote on each topic.